Simulation of a Trust and Reputation based Mitigation Protocol for a Black Hole Style Attack on VANETs

Agostino Martorana
Department of Computer Science
Middlesex University London
United Kingdom
Email: am2847@live.mdx.ac.uk

Giuseppe Primiero
Department of Computer Science
Middlesex University London
United Kingdom
Email: G.Primiero@mdx.ac.uk

Jacopo Tagliabue
Tooso Inc.
United States
Email: jacopo.tagliabue@tooso.ai

***Abstract*—From a security standpoint, VANETs (Vehicular ad hoc Networks) are vulnerable to attacks by malicious users, due to the decentralized and open nature of the wireless system. For many of these kinds of attacks detection is unfeasible, thus making it hard to produce security. Despite their characterization as dynamically reconfigurable networks, it is nonetheless essential to identify topology and population properties that can optimise mitigation protocols' deployment. In this paper, we provide an algorithmic definition and simulation of a trust and mitigation based protocol to contain a Black Hole style attack on a VANET. We experimentally show its optimal working conditions: total connectivity, followed by a random network; connection to external networks; early deployment of the protocol and ranking of the message. We compare results with those of existing protocols and future work shall focus on repeated broadcasting, opportunistic message forwarding and testing on real data.**

1. Introduction

MANET (Mobile Ad hoc Network) refers to a self-configuring system of mobile routers, with wireless links to form an arbitrary topology. The mobility of the routers are provided randomly and organized arbitrarily. VANET (Vehicular Ad hoc Network) is the application of MANET structures to vehicles and roadside unit networks created to enhance transportation systems through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. This kind of systems has various potential applications [10], [21]. From a security standpoint, any MANET is highly vulnerable to attacks by malicious users and security of data transmission is the main concern. Due to their distributed and dynamic nature, such networks are open to several types of threats, including false message propagation. For many of these kinds of attacks, detection is unfeasible, making it hard to produce security. Trust and reputation are among the most used concepts to ensure integrity, reliability and safety of services. While an increasing literature is available on mitigation protocols for such attacks, the focus on dynamic, reconfigurable networks tends to make these results opaque with respect to conditions of the network when the attack takes place. It is nonetheless essential to identify topology and population properties that can optimise mitigation protocols' deployment on VANETs.

The present paper simulates a type of attacks on a VANET and deploys a trust and reputation model to mitigate it presented in [17]. Our main aim is to investigate whether this protocol is effective in preventing such type of attacks under precise and well specified conditions. The algorithms that constitute the backbone of the implementation are presented as pseudo-code below (see Figures 1,3,5,6), and the code of our simulation is made available at https://github.com/gprimiero/trust4vanet2 for reproducibility purposes, together with all data from our experiments. Contrary to many contributions in the area, we make use of standard network theory analysis to investigate which network properties (like size and topology), which population properties (like proportion of the attackers) and which contextual conditions (like the current content distribution) are optimal to constraint the attack. This allows to set some clear benchmarks for the specific type of attack we consider; moreover, it makes the model general enough to be easily reconfigured for V2I scenarios. The rest of the paper is structured as follows: in Section 2 we overview first the types of attacks on mobile networks defined in the literature and some of the current protocols deployed for their mitigation; in Section 3 we illustrate informally the kind of attack object of our analysis and in Section 4 we provide an algorithmic description in pseudo-code of our trust and reputation-based mitigation protocol; in Section 5 we present our experimental results, offer some comparisons with existing data and highlight advantages and limitations of our approach; finally, in Section 6 we briefly present further steps of this research.

2. Related Work

Due to their particular architecture, ad-hoc networks are more easily attacked than wired networks. Two main kinds of attacks are usually distinguished [11, p.956]:

- passive attacks: they do not disrupt the operation of the protocol, but try to discover valuable information by listening to traffic;

- active attacks: they inject arbitrary packets and try to disrupt the operation of the protocol in order to limit availability, gain authentication, or attract packets destined to other nodes.

We focus below on active attacks. The most common type, performed due to their easiness, are "Attacks Using Modification": a malicious node disturbs the good operation of an ad-hoc network by announcing better routes (to reach other nodes or just a specific one) than the other nodes. This kind of attack is based on the modification of the metric value for a route, or executed by altering control message fields [11, p.956]. In "Attacks using Impersonation" (spoofing), the malicious node hides its real IP address or MAC address and uses another one. As current ad-hoc routing protocols like AODV and DSR do not authenticate source IP address, a malicious node can launch many attacks by using spoofing. For example, a hacker can create loops in the network to isolate a node from the remainder of the network. To do this, the hacker just has to take the IP address of another node in the network and then use it to announce a new route (with the smallest metric) to the others nodes. By doing this, he can easily modify the network topology as he wants [11, p.956]. Attacks by "Resource Overuse" indicate the additional use of a resource by a node for any activity other than route finding and maintaining or transmitting data. As mobile nodes may have limited resources in terms of memory, storage, processing power, and battery life, if an IDS involves too much data and computations, then there will be more usage of memory and processing power, thus disrupting the normal functioning of the network [9, p.388].

Several attack strategies can be identified [5, p.345]:

- **Black hole Attack**: in this type of attack, malicious nodes broadcast the message to all the nodes, diverting all the traffic toward themselves, and without forwarding the data packets to the neighbouring nodes, so that all the (non-malicious) data packets are dropped;
- **Gray hole Attack**: in this form of Black hole attack, the malicious nodes drop the data packets for particular nodes for particular period of time in the network. Gray hole attacks are more difficult to identify as compared to black hole attacks;
- **Wormhole Attack**: in this type of attack, two malicious nodes from a tunnel and all the data packets received at one location of the network are tunneled to the other location, in such a way that all the data are resent to the network. The tunnel between two malicious nodes is called a Wormhole. Such attacks prevent any route other than through the wormhole from being discovered;
- **Byzantine Attack**: this type of attack is carried out by intermediate nodes or group of intermediate nodes. Such malicious nodes provide the false routing information and create routing loops as well as forward their data packets to that path which is not optimal and which may be harmful to the system;
- **Denial of Service Attack**: it prevents the victim from using all or part of the network connections. DOS attack may have numerous forms and are hard to detect. In this type of attack, attacker nodes send the excessive amount of data packets or requests to the server so the latter gets busy in testing illegal request and will not be available to other. This attack may degrade the performance of the network since it consumes the energy (Battery Power) of nodes.

Black hole detection has been an active area of research since the 'next hop information' based scheme was proposed in 2002 [6], with a few among the proposed solutions focusing on collaborative black holes. Among the most sophisticated mitigation protocols is the 'Fidelity Table' method [20]. In this model, every participating node is allotted a particular fidelity level, a measure of reliability. Whenever a source node broadcasts a RREQ (Route Request) and holds up, the incoming RREPs (Route Reply) are gathered in its Response Table. If the average of the fidelity level of RREP sending node (RREPN) and its next hop node (NHN) in the route is found to be over a predetermined threshold, the RREPN (Route Reply Sending Node) is considered as trustworthy. Therefore, on the receipt of multiple RREPNs, the one with the highest fidelity level is selected. However, if multiple nodes have the same fidelity level, the RREPN with the minimal hop count is chosen. Finally, routing is accomplished via the selected path. Upon data receipt, the destination node sends an acknowledgement to the source node within timer. Next, fidelity level of the RREPN is incremented as an accolade for honest routing else that of both RREPN and its NHN is decremented for being collaborative. If fidelity level of a node drops to zero, it is considered as a black hole and the presence of attack is intimated to all using alarm packets. Despite the fact that this method handles both single and collaborative black hole attacks, it involves increased storage overhead, routing overhead, computational overhead and delay. This is because each node should maintain a Fidelity Table and a Response Table that must be updated and exchanged among the nodes periodically [19]. Reputation is a crucial notion in several other protocols, and used in ours as well.

Reputation is complemented by trust models in several VANET models, defined in accordance to their main object: entity-centric [13], [8], data-centric [18], [12] and combined [23]. An overview of the issues related to trust in fixed and mobile ad hoc networks is given in [24], while approaches for trustworthiness and reputation in ad hoc mobile networks are presented, for example, in [7], [4]. The work in [22] offers an analysis that accounts for reputation as a characteristic of message forwarding among vehicles, drivers and other agents: reputation of these agents is based on a descriptive ontology and is used to provide feedback in the system. In [17], we have provided a logic calculus that formally verifies this reputation model by adding a trust function in order to guarantee absence of unsafe behaviours. An interesting way to test the reliability of the model is to simulate an attack, which is the main aim of the present

```
to discovery [ _propositionIndex _# turtles ]
ask n−of _#turtles turtles [ knowProposition _propositionIndex −1 1 ]
end
```

Figure 1. Discovery

contribution.

3. The Attack Scenario

In view of the taxonomy of attacks on mobile networks presented above, we restrict our current attention specifically to a Black Hole type of attack in which one or more nodes block the transmission of truthful information and start distributing manipulated data. For the present purposes, it is irrelevant the content or dimension of this data and we just represent it by atomic formulas and their negation: an advanced modeling of the present protocol will focus on the use of real data.

At any point in time in which the attack is deployed, the network can be assumed to have a specific topological configuration produced by the actual set of active communications among its nodes (see Section 4.1 for the topologies analysed in the present work). A first aim of the present work is to establish which topological structure should the agent have at the moment of the attack (or immediately after) in order to maximally constrain it.

Agents (vehicles) are categorised according to one of three families (or breeds, the labels used are ours, but they reflect standard typologies of agents in similar scenarios):

- *discoverers*: one or more agents in possession of truthful and updated data, possibly received by an external network through Internet or a RSU;
- *attackers*: one or more agents performing the attack, by flipping the truth value of data to be distributed in the network;
- *receivers*: the remaining set of vehicles, which have no information, or with information requiring an update.

Discoverers generate truthful information by assigning truth-value 1 to a proposition which describe some content of relevance to the network (e.g. "temperature $< 5c$"), according to the routine defined in Figure 1.

Message passing is performed by the routine in Figure 2. It requires every agent with a message whose truth value is non-neutral (i.e. equal to either 1 or 0) to pass it to one of its neighbors. Note that currently we do not implement the opportunistic forwarding protocol described in [22], which requires to select the recipient with the highest reputation value.

By definition, this routine holds for both discoverers and attackers. When an attacker receives truthful information p (labeled by truth-value 1) on the current outside temperature and road conditions in view of ice and snow, its aim is to flood the network with its negation (labeled by truth-vale 0), and to get as many vehicles as possible to distribute it. We

```
to spreadKnowledge [ _propositionIndex ]
ask turtles with [ item _propositionIndex propositions != 0 ]
[
let myId who
let truthVal item _propositionIndex propositions
ask one−of link−neighbors
[ knowProposition _propositionIndex myId truthVal ]
]
end
```

Figure 2. Message Passing

```
to attack [ _propositionIndex ]
ask n−of number_attacker turtles with
[ item _propositionIndex propositions = 1 ]
[
set breed attackers
knowProposition _propositionIndex −2 −1
]
end
```

Figure 3. Attack

implement the attack through the routine defined in Figure 3.

If no protocol is in place to resolve conflicting information transmission, the attack might easily succeed. This is clearly shown by the initial experimental results in Section 5. If a simple majority protocol is used to discern between conflicting data, it might be enough for the attacker to target a limited number of nodes with high degree of connectivity, in order to flood the entire network. Trust and reputation are crucial to improve on simple quorum protocols.

Several of the work presented in Section 2, introducing trust and reputation methods for VANET attacks, presents such attacks under the general assumption that the network is dynamic and therefore no topological analysis is usually done. As a negative effect, this means that results are usually under-specified with respect to which are the initial conditions (topological, temporal, etc.) of the attack. We investigate the preferred topological structure of the network in the aftermath of an attack in order to limit its efficacy. In particular, our parameters concern:

1) the topology of the network at the moment of the attack;
2) the proportion of discoverers and attackers;
3) the ranking of truthful information object of the attack;
4) the state of truthful information diffusion at the moment of the attack.

4. The Trust and Reputation Model

In the present section we provide a high-level description of our trust and reputation based model, deployed to mitigate Black-Hole attacks on mobile networks. This protocol is based on the logic $(\mathtt{un})\mathtt{SecureND}$, a natural deduction calculus introduced in [16] to define trust, mistrust and distrust protocols and extended in [15] with a negation connective. In [17] a modified version has been presented, adapted for a VANET network with the introduction of a reputation

```
to-report calculateApperceptionValue
[ sourceRanking _propositionRanking _tick ]
report (1.0*sourceRanking)*(1.0*_propositionRanking)*( _tick *1.0)
end
```

Figure 4. Apperception

measure. This last version is implemented in NetLogo in this work for the purpose of experimental analysis.

Given $\mathcal{A}$ a set of agents containing vehicles $\mathcal{V}$ and roadside units (RSUs) $\mathcal{R}$, an order $\prec$ holds between agents and expresses a reputation order. $\mathcal{S}$ denotes a set of services for the messages, with $\mathcal{C}$ a set of service characteristics, and each element $C^{S_i}_{\overrightarrow{n}}$ denoting the set of n characteristics of service S_i. In general, characteristics $C^{S_i}_{\overrightarrow{n}}$ of services for each service S_i are associated with an order $\leq$ used to order messages, and for two characteristics $C^{S_i}_{\overrightarrow{n}}, C^{S_j}_{\overrightarrow{n}}, i \neq j$, there is no order between them. In the present experimental analysis, we limit ourselves to a single message p about a given service with one characteristic. The analysis of the model generalised to multiple messages concerning several characteristics for one service is left to future research.

In the implementation, we translate an order relation $p_i < p_j$ between messages as a fixed absolute ranking value of a unique message p (`rank p`), which we will assume is shared by all vehicles in each messaging instance, and can assume three distinct numerical values corresponding to low, medium and high relevance. As a further simplification, we assume atomic messages, although the underlying logic (un)`SecureND` allows for closure under connectives. For any given message p received from another agent (either vehicle or RSU), a vehicle will collect all the formulas that follow from accepting it. This is called the Feedback Set of an agent with respect to a message. In view of the mentioned simplification, our Feedback Set is always a singleton. The vehicle assigns a value to this atomic formula, using three parameters:

1) a fixed ranking of the source, generated in this model automatically by the topology of the network (see below Section 4.1 for more details);
2) a ranking of the message, fixed at some low, medium or high value for each messaging operation;
3) the time at which the reception of the message occurs, computed by the clock underlying the simulation (the later the message arrives, the more updated it is considered, the highest its value).

We call the resulting value `Apperception`, see Figure 4. As a result of working with atomic messages only, the Apperception of a vehicle for a message is computed by unary factors. Using this value, we define directly the order of reputation for agents, which establishes a higher position for the vehicle whose apperception is greater.

4.1. Network Construction

As mentioned, the original model which inspires the present implementation and presented in [22] uses reputation to define a recipient selection protocol: after v_i broadcasts a 'hello' message, if both v_k, v_j receive and accept the message, then v_i has to select a recipient on the basis of the reputation order between v_k and v_j. Accordingly, a new profile is built out of v_i and the higher of the two recipients, thus modeling a communication channel.

In the present implementation, given fixed network topologies we assign an agent ranking on their basis. Reputation then is defined by this factor together with message ranking and timing, as explained above. In particular, we consider the three following main topologies and associated agent ranking methods:

- *small-world*: the network is generated according to a power-law, by which each new node has a higher probability to be linked to a node with high in-degree than to one with low in-degree; in this topology, higher in-degree nodes have higher ranking;
- *total*: every node is connected to any other node, and all have the same ranking;
- *random*: edges are randomly distributed, and so is the ranking which reflects the in-degree of each node.

The advantage of fixing the networks and of determining the agent ranking on that basis is to experimentally evaluate the optimal topological conditions for the mitigation protocol. A next step of this protocol evaluation will implement reputation (based on message ranking and timing only) as a criterion for recipient selection.

4.2. Message Passing Protocol

The trust protocol from the logic (un)`SecureND` enforces a consistency check on message passing. Each valid vehicle profile meets all the requirements and conflicts clauses of all service messages that the vehicle receives. Trust allows to select valid messages among those that are read: if a message is received by a vehicle and it preserves its profile consistency, then it can be trusted. A message readable and trusted by a vehicle can be further broadcast to other vehicles. Mistrust is the protocol acting on conflicting messages: a currently held message conflicting with a newly arrived message is removed from the current vehicle profile and none of its consequences are included; any message consistent with the conflict resolution is trusted by removal of the mistrusted message in the vehicle profile, including any required dependency.

In the present implementation, message passing with consistent information or when the receiving agent is an attacker is unproblematic: it induces respectively acceptance and rejection of the new message (in the latter case, it means an attacker once flipped the value of a received message will no longer change it). Message passing consistent information to an agent who has no prior information induces a reputation control: information is accepted when coming from a source with higher reputation, while the receiver remains undecided if information comes from a source with lower reputation. A conflict is generated by two

```
to-report verifyProposition
[ _currentVal _newVal _propositionIndex _breed _currentWho _sourceWho]
if (_breed = attackers )
[ report _currentVal ]
if ( _currentVal = _newVal)
[ report _newVal ]
if ( _currentVal = 0)
[ ifelse (sourceApperception >= currentApperception)[ report _newVal ]
[ report 0 ]
]
if ( _currentVal != _newVal)
[ ifelse (sourceApperception >= currentApperception)
[ let consensus
neighborConsensus _propositionIndex _newVal sourceApperception
ifelse (consensus = false ) [ report _currentVal ]
[ report _newVal ]
]
[ report _currentVal ]
]
```

Figure 5. Trust

```
to-report neighborConsensus
[ _propositionIndex _truthVal _sourceApperception ]

let total_neighbors_with_opinion count link-neighbors with
[item _propositionIndex propositions != 0
and item _propositionIndex apperception > _sourceApperception]

let total_neighbors_with_same_opinion count link-neighbors with
[item _propositionIndex propositions = _truthVal
and item _propositionIndex apperception > _sourceApperception ]

if ( total_neighbors_with_opinion = 0) [ report false ]
report total_neighbors_with_opinion = total_neighbors_with_same_opinion
end
```

Figure 6. Consensus

contradictory messages, and the profile is valid when such conflicts are avoided. Receiving contradictory information requires again reputation checking: with a source that has reputation at least as high as the receiver, the latter checks with any linked agent with higher reputation than the source; when these show consensus (see Figure 6) with the receiver about its currently held information, the latter does not change its mind, and it does otherwise. See Figure 5.

5. Experimental Results

The model presented above has been implemented in NetLogo for experimental validation and analysis. With this aim, our focus has been on the following parameters:

1) *network topology*: we investigate the optimal topology among small-world, total and random networks for the minimization of the effects of the attack;
2) *network size*: we consider three main values on this parameter, with networks of $10, 50, 100$ nodes; these sizes are only indicative of small, medium and large networks, and analysis on sizes lower than 10 and greater than 100 agents have shown no statistically significant differences in the results;
3) *message ranking*: a static value at each message passing operation, valid for all agents (both discoverers and attackers), investigated with three different values $0.20, 0.50, 0.80$ corresponding respectively to `low`, `medium`, `high` ranking;
4) *proportion of discoverers*: proportion of agents with truthful information at the beginning of the protocol, investigated with three values $10\%, 50\%, 90\%$; we use these values as indicative of networks with respectively a low, medium and high level of external connectivity (e.g. through road-side unit connected to the Internet);
5) *proportion of attackers*: proportion of agents performing an attack, investigated with two values $1\%, 10\%$; these values are experimentally shown to be the lower and higher bound required for optimization: with the lower bound we have a minimum of 1 attacker, which is a standard configuration in several other experimental analyses; with the higher bound we establish that our protocol guarantees that results do not get worse when more than 10% of the population works as an attacker;
6) *network coverage for attack*: coverage of the entire network with truthful information when the attack is executed, investigated with three values $20\%, 50\%, 80\%$ corresponding respectively to `low`, `medium`, `high` coverage (reflecting an early, intermediate and late attack).

Note that given the above configurations, our best scenario will present 1% of attackers and 90% of discoverers, while the worst one will have 10% for both categories of agents: we stress here that experiments on intermediate configurations fall in the appropriate range thus confirming these to be the most interesting cases.

All experiments reported below have been executed on a machine with 7.7 GB of memory, 64bit Ubuntu 17.10 system, NetLogo 6.0. Each configuration has been run for a minimum of 100 times, and for several configurations we have run 10 repetitions of 100 runs each. The results report the median values of all runs.

5.1. First Configuration: Topology and Size

In the first set of experiments, we focus on determining optimal topology and size for our mitigation protocol. We investigate all sizes and topologies while keeping message ranking and network coverage parameters both at a minimum level (respectively 0.20 and 20%).

Our first experiment is executed with a proportion of attackers and of discoverers both fixed at 10%, i.e. with respectively $1, 5, 10$ attackers/discoverers on networks of $10, 50, 100$ nodes. In this way we also replicate attacks performed by a single and by multiple agents. Results are plotted in Figure 7: it clearly illustrates how under these conditions the attack is highly successful and the protocol has limited efficacy. The plotted lines reflect the three topologies: under these attack conditions, the infection remains high over all three and across all sizes.

These values of infection spread tend to decrease sensibly when varying the proportion between attackers and discoverers, in a way which preserves in general the optimality of the topologies. By fixing the proportion of attackers at

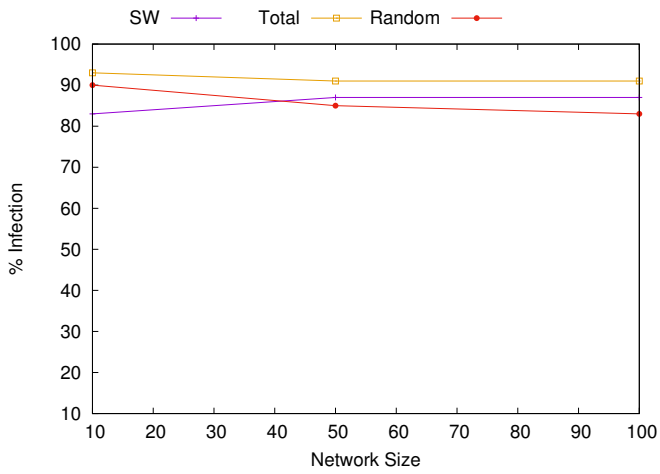

Figure 7. Experiment 1: Attack with 10% attackers/discoverers.

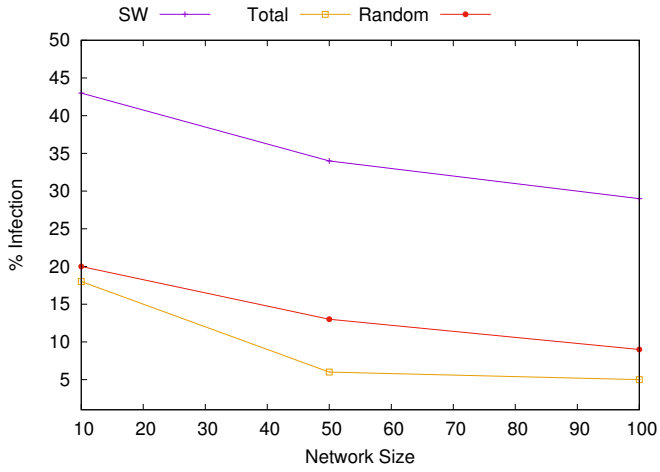

Figure 8. Experiment 2: Attack with 1% attackers and 90% discoverers.

1% (in general, this means we are constraining the analysis to a single attacker scenario) and increasing the proportion of discoverers to 90% (such increase should be understood as a large majority of the network being connected to an external information source, like RSUs), the protocol makes the network a lot more resilient to the attack, see Figure 8. These results shows an immediate sensible dropping of the infection values, with total networks being the best performing, followed by random and small-world ones, and with larger networks being less prone to infection than smaller ones.

Performed – but not reported here – experiments cover several in-between configurations: in general, networks of size < 10 nodes do not offer improvements and networks of size > 100 do not present worse results, confirming our to be maximally and minimally optimal configurations relative to size and attackers/discoverers proportions.

These initial findings show that, while there is a positive effect intrinsically generated by the network size (the greater the network, the harder to spread the attack), it is

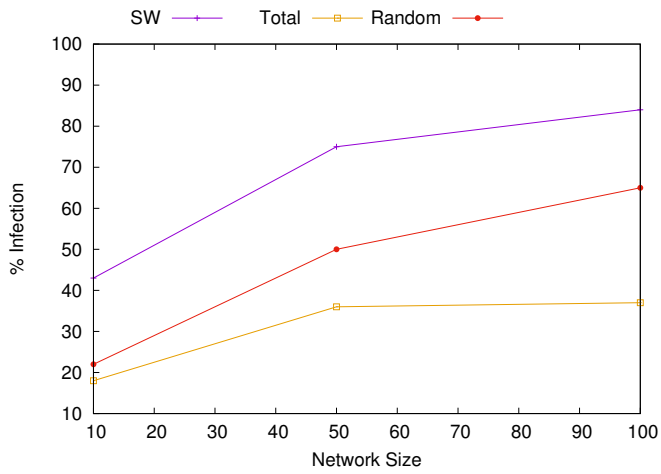

Figure 9. Experiment 3: Attack with 10% attackers/discoverers and 0.5 message ranking.

sufficient for the attackers breed to cover 10% of the entire population in order to nullify such advantage. On the other hand, with a single attacker, size sensibly helps reducing the infection, with a totally connected network presenting the most advantageous setting for the mitigation protocol, followed by random and small-world. This suggests that in a dynamically reconfigurable network, if an attack is identified or suspected, total connectivity should be sought by the agents (vehicles and RSU) in order to minimize its negative effects.

5.2. Second Configuration: Message Ranking

The second set of experiments focuses on the variable `rankp`, a numerical value to express a ranking of the message: the higher this value, the more relevant is the message considered by the network. We investigate the above mentioned three values: $0.2, 0.5, 0.8$ to express low, medium and high relevance. The experiments are meant to show how such ranking affects the ability of the protocol to constraint the attack. Rising the value of `rankp` from 0.2 to 0.5 in the setting of the first experiment (Figure 7) has an immediate positive effect, bringing a drastic improvement in small networks of 10 nodes, and a significant one also in larger configurations, see Figure 9. Further rising the value of `rankp` to 0.8 does not improve these results.

On the other hand, analysing small-world, total and random networks with a single attacker and a proportion of 90% discoverers (i.e. the configuration of Experiment 2 in Figure 8), rising message ranking has no influence in further reducing the infection in all topologies: the results confirm the difference between the three topologies reflected by Experiment 2, with total network showing the least infection, an average of 50% less than in random networks, which in turn on average is almost 60% better than small-world networks. These results suggest that reputation is only partially influenced by the relevance of the message, as this obviously is computed for both discoverers and

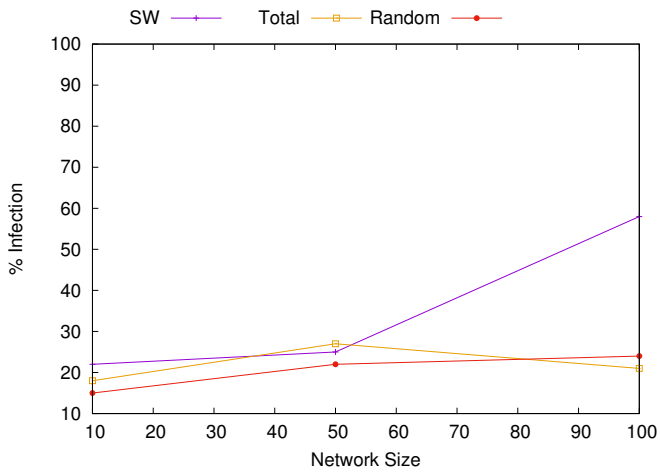

Figure 10. Experiment 4: Attack with 10% attackers/discoverers and 0.8 network coverage.

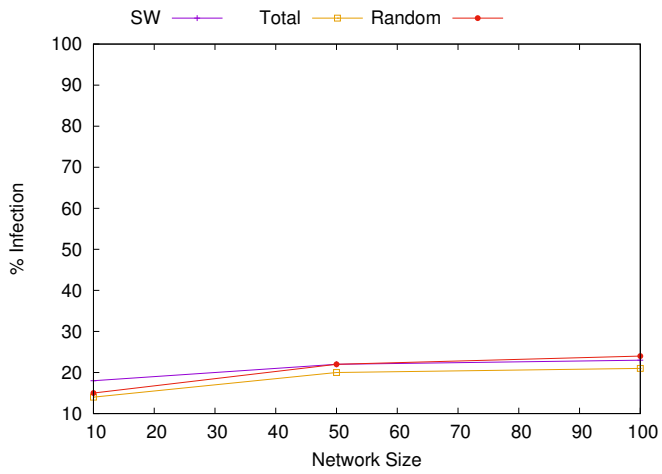

Figure 11. Experiment 5: Attack with 10% attackers/discoverers, 0.5 message ranking and 0.8 network coverage.

attackers. Future experiments should focus on the difference to this factor provided by messages concerning different characteristics within the same service.

5.3. Third Configuration: Network Coverage

The third set of experiments focuses on the variable `network coverage`, a numerical value to express how much of the overall network is informed of the correct message when the attack is struck: the lower this value, the less the network is correctly informed when the attack begins. We investigate the above mentioned three values: $0.2, 0.5, 0.8$ to express low, medium and high coverage. Our aim is to investigate how the infection is effected when parametrised by the relation between the attack being struck and a percentage of the network having already received truthful data. We show in Figure 10 the results for the setting of the first experiment (Figure 7) modified by rising the value of `network_coverage` from 0.2 to 0.8: in this setting a drastic improvement in small networks of 10 nodes is obtained, with all topologies managing to constrain the infection to a maximum of 25%; in larger configurations this level is maintained, except for small-world networks presenting a pick of over 50% infection with size of 100 nodes. In general this means that a large fraction of the network truthfully informed is essential to constrain infection.

Note that, again, analysing small-world, total and random networks with a single attacker and a proportion of 90% discoverers (i.e. the configuration of Experiment 2 in Figure 8), while rising the level of the variable `network_coverage` from 0.2 to 0.8, no significant improvement in the results is obtained. The difference between the three topologies remain invariant, with on average total networks showing 50% less infection than random and these 50% less than small-world (these differences are slightly less for small networks of 10 nodes). These results suggest that damage limitation after an attack of the present type requires an elevated number of nodes connected to an external network, to facilitate the increase of agents with updated information at any point.

5.4. Fourth Configuration: Combined Optimal Parameters

In the last set of experiments, we combine the optimal message ranking and network coverage parameters. We first observe this combination in the non-optimal setting of experiment 1 (Figure 7), i.e. with a proportion of attackers and of discoverers both fixed at 10% on networks of $10, 50, 100$ nodes. The results are plotted in Figure 11: this configuration manages to keep the infection below 25% in the largest networks and presents an average decrease in infection of over 75% across the three topologies, when compared with the worst-case scenario of experiment 1.

Finally we consider how the optimal configuration of message ranking and network coverages affects the results of Experiment 2 (Figure 8), where the proportion of attackers is minimised to 1% and that of discovered maximised to 90%. Results shown in Figure 12 illustrate how the protocol is able to minimise the negative impact of an attack by a single agent on networks with a large proportion of agents who transmit the truthful data if messaging ranking and network coverage are optimised. The average improvement on the result on the same networks without these last two parameters optimised is of almost 30% across all topologies.

5.5. Summary of Results, Comparisons and Limitations

Our experimental results can be summarised as follows:

- the reputation protocol is essential in constraining the attack on networks of any size with at least 10% of the population acting as attackers, presenting an improvement of up to 80% on the same conditions without reputation;

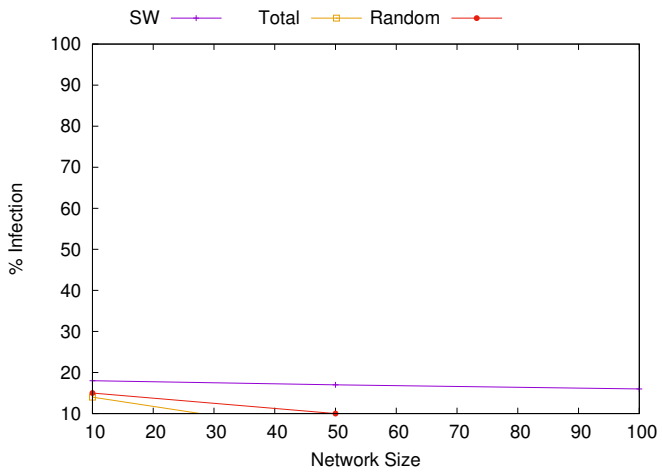

Figure 12. Experiment 6: Attack with 1% attackers, 90% discoverers, 0.5 message ranking and 0.8 network coverage.

- the protocol is up to 90% more efficient in a one-attacker condition when compared to attacks executed by up to 10 malicious agents;
- total networks are overall the most efficient topology for attack mitigation, with results up to 15% better than in random and 30% better than in small-world networks;
- network size is relevant only when reputation takes into account a higher ranking for the message: then the difference between infection in a 10 nodes network and in a 100 nodes one can reach up to 30%.

These initial results suggest that the optimal deployment strategy for the present mitigation protocol requires that:

1) agents seek total connectivity, and while this is an unlikely configuration in a real-world scenario, a random distribution of edges performs better than a power-law distribution and it is sufficiently close to the optimal topology;
2) the protocol is deployed as soon as possible, hence ideally from the start of the messaging operation rather than as a reaction to the attack, so that the number of agents with correct information is as high as possible when the attack occurs;
3) the network is as much as possible (in terms of number of agents) connected to external information sources.

Other positive aspects of the protocol are that its results are not affected by either the size of the network, or by a number of attackers greater than 10% of the entire population.

To offer some comparison with other protocols, in the following we relate the infection spread contained by our protocol with the packet delivery ratio in other protocols, i.e. the ratio between the number of packets sent by the sources and the number of packets received by the sink at the final destination. The two values are obviously inversely proportional, although the comparison is only partially significant given the different topological structures and other incomparable parameters. The first comparison is made with the protocol from [20] on a 25 nodes network with 2 attackers: results reported for PCBHA with random way point model mobility, 100 items load and $5-8$ transactions, give a delivery ratio of 60%, an increase of around 90% compared with the AODV protocol from [3]; the protocol based on (un)`SecureND` with random network of 25 nodes, 2 attackers and 23 discoverers (note the difference with the 100 packets load), 0.5 `rankp` and 0.8 `network_coverage`), guarantees an average infection spread limited to 17% over 1000 simulation runs; considering the initial 8% infection due to the attackers, the comparable delivery packet ratio amounts to 91%.

We provide a second comparison with the results from [2], where the protocol TCRSR on a 50 nodes network with 10% attackers is reported to guarantee up to 50% delivery ratio when agents move at 10m/s speed; on the same network size and with the same percentage of attackers, our protocol with random network of 50 nodes, 5 attackers and 45 discoverers, 0.5 `rankp` and 0.8 `network_coverage`), guarantees an average infection spread limited to 22% over 1000 simulation runs; considering the initial 10% infection due to the attackers, the comparable delivery packet ratio amounts to 88%.

It is also essential to stress some limitations of our method. First of all, the high-level of abstraction proper of our implementation does not allow to analyse low-level properties, like packet size. Although possible, we have not implemented other useful properties, like number of parallel communications and communication distance (currently we allow message passing by physical presence on the same world's patch, roughly corresponding to close contact of the agents). Secondly, the current experiments are performed in an idealised setting, where additional problems like physical obstacles or weaknesses in the communication protocol are not taken into account: only some physical implementation could further enhance our results by taking these aspects into account.

6. Conclusions

In this paper we have provided a simulated analysis of a trust and reputation based protocol for the mitigation of Black Hole type attacks in VANETs. The present implementation is simplified under several aspects, in particular: messaging is a one-time event and is not repeated in short time spans; message forwarding happens by random recipient selection; messages are atomic. Notwithstanding this specification, our analysis clearly shows a working protocol which is especially efficient with total and random network, if deployed early and with a large connectivity to external information sources. Next steps of this research include: the implementation of a lower level of abstraction in the protocol, repeated broadcasting and opportunistic forwarding; the translation to a simulation environment that can help capture such relevant properties, like OmNet++ or VSimRTI; testing the protocol on real data, which will be made available

thanks to the testbed deployed by colleagues at Middlesex University [14].

References

[1] *2017 IEEE European Symposium on Security and Privacy Workshops, EuroS&P Workshops 2017, Paris, France, April 26-28, 2017*. IEEE, 2017.

[2] Rajkumat Banoth and G. Narsimha. Trust based certificate revocation for secure routing in manet. *Procedia Computer Science*, 92:431441, 2016.

[3] S. Das C.E. Perkins, E. Belding-Royer. Ad-hoc On Demand Distance Vector (AODV) Routing. Technical report, Network Working Group RFC 3561, July 2003.

[4] Brijesh Kumar Chaurasia, Ranjeet Singh Tomar, and Shekhar Verma. Using trust for lightweight communication in vanets. *IJAISC*, 5(2):105–116, 2015.

[5] Nilesh N. Dangare and R. S. Mangrulkar. Design and implementation of trust based approach to mitigate various attacks in mobile ad hoc network. *Procedia Computer Science*, 78(1):554–561, 2016.

[6] Hongmei Deng, Wei Li, and Dharma P. Agrawal. Routing security in wireless ad hoc networks. *Communications Magazine*, pages 70–75, 2002.

[7] John Finnson, Jie Zhang, Thomas T. Tran, Umar Farooq Minhas, and Robin Cohen. A framework for modeling trustworthiness of users in mobile vehicular ad-hoc networks and its validation through simulated traffic flow. In Judith Masthoff, Bamshad Mobasher, Michel C. Desmarais, and Roger Nkambou, editors, *User Modeling, Adaptation, and Personalization - 20th International Conference, UMAP 2012, Montreal, Canada, July 16-20, 2012. Proceedings*, volume 7379 of *Lecture Notes in Computer Science*, pages 76–87. Springer, 2012.

[8] Félix Gómez Mármol and Gregorio Martínez Pérez. Trip, a trust and reputation infrastructure-based proposal for vehicular ad hoc networks. *J. Netw. Comput. Appl.*, 35(3):934–941, May 2012.

[9] Muhammad Imran, Farrukh Aslam Khan, Tauseef Jamal, and Muhammad Hanif Durad. Analysis of detection features for wormhole attacks in manets. *Procedia Computer Science*, 56:384 – 390, 2015. The 10th International Conference on Future Networks and Communications (FNC 2015) / The 12th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2015) Affiliated Workshops.

[10] Vinita Jindal and Punam Bedi. Vehicular ad-hoc networks: Introduction, standards, routing protocols and challenges. *IJCSI International Journal of Computer Science Issues*, 13(2):1694–0784, 2016.

[11] Praveen Joshi. Security issues in routing protocols in manets at network layer. *Procedia Computer Science*, 3(1):954–960, 2011.

[12] Nai-Wei Lo and Hsiao-Chien Tsai. A reputation system for traffic safety event on vehicular ad hoc networks. *EURASIP J. Wirel. Commun. Netw.*, 2009:9:1–9:2, April 2009.

[13] U. F. Minhas, Jie Zhang, T. Tran, and R. Cohen. A multifaceted approach to modeling agent trust for effective communication in the application of mobile ad hoc vehicular networks. *Trans. Sys. Man Cyber Part C*, 41(3):407–420, May 2011.

[14] Vishnu Vardhan Paranthaman, Arindam Ghosh, Glenford Mapp, Victor Iniovosa, Purav Shah, Huan X. Nguyen, Orhan Gemikonakli, and Shahedur Rahman. Building a prototype vanet testbed to explore communication dynamics in highly mobile environments. In Song Guo, Guiyi Wei, Yang Xiang, Xiaodong Lin, and Pascal Lorenz, editors, *Testbeds and Research Infrastructures for the Development of Networks and Communities*, pages 81–90, Cham, 2017. Springer International Publishing.

[15] Giuseppe Primiero. A calculus for distrust and mistrust. In Sheikh Mahbub Habib, Julita Vassileva, Sjouke Mauw, and Max Mühlhäuser, editors, *Trust Management X - 10th IFIP WG 11.11 International Conference, IFIPTM 2016, Darmstadt, Germany, July 18-22, 2016, Proceedings*, volume 473 of *IFIP Advances in Information and Communication Technology*, pages 183–190. Springer, 2016.

[16] Giuseppe Primiero and Franco Raimondi. A typed natural deduction calculus to reason about secure trust. In Ali Miri, Urs Hengartner, Nen-Fu Huang, Audun Jøsang, and Joaquín García-Alfaro, editors, *2014 Twelfth Annual International Conference on Privacy, Security and Trust, Toronto, ON, Canada, July 23-24, 2014*, pages 379–382. IEEE Computer Society, 2014.

[17] Giuseppe Primiero, Franco Raimondi, Taolue Chen, and Rajagopal Nagarajan. A proof-theoretic trust and reputation model for VANET. In *2017 IEEE European Symposium on Security and Privacy Workshops, EuroS&P Workshops 2017, Paris, France, April 26-28, 2017* [1], pages 146–152.

[18] Maxim Raya, Panagiotis Papadimitratos, Virgil D. Gligor, and Jean-Pierre Hubaux. On data-centric trust establishment in ephemeral ad hoc networks. In *INFOCOM 2008. 27th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 13-18 April 2008, Phoenix, AZ, USA*, pages 1238–1246. IEEE, 2008.

[19] Arathy K Sa and Sminesh C Na. A novel approach for detection of single and collaborative black hole attacks in manet. *Procedia Computer Science*, pages 264–271, 2016.

[20] Latha Tamilselvan and V. Sankaranarayanan. Prevention of co-operative black hole attack in manet. *Journal of networks 3.5*, pages 13–20, 2008.

[21] Ravi Tomar, Manish Prateek, and G. H. Sastry. Vehicular adhoc network (vanet) - an introduction. *International Journal of Control Theory and Applications*, 9(18):8883–8888, 2016.

[22] R Vanni, L.M.S. Jaimes, G. Mapp, and E. Moreira. Ontology driven reputation model for vanet. In *AICT 2016, The Twelfth Advanced International Conference on Telecommunications, IARIA*, pages 14–19, 2016.

[23] Yu-Chih Wei and Yi-Ming Chen. *Reliability and Efficiency Improvement for Trust Management Model in VANETs*, pages 105–112. Springer Netherlands, Dordrecht, 2012.

[24] Philipp Wex, Jochen Breuer, Albert Held, Tim Leinmüller, and Luca Delgrossi. Trust issues for vehicular ad hoc networks. In *Proceedings of the 67th IEEE Vehicular Technology Conference, VTC Spring 2008, 11-14 May 2008, Singapore*, pages 2800–2804. IEEE, 2008.